

Exhibit A – Document Production Format

1. Definitions

The following definitions apply to these instructions:

A. **“Database”** means an electronic collection of structured data (often maintained in a non-custodial manner), such as data created and maintained in Oracle, SAP, SQL, or Microsoft Access.

B. **“ESI” or “Electronic Document”** refers to information stored in electronic form including word processing files (*e.g.*, Microsoft Word), computer presentations (*e.g.*, Microsoft PowerPoint), spreadsheets (*e.g.*, Microsoft Excel) and email, together with the available metadata associated with each such document.

C. **“Extracted Text”** refers the result of the process by which textual content of an Electronic Document is extracted by eDiscovery processing or indexing software.

D. **“Load File”** refers to a file or files issued with each production providing a map to the images and metadata or coding associated with the records in the production.

E. **“Native Format”** as used herein, means the default format of an Electronic Document created by its associated software program. For example, Microsoft Excel produces its output as ‘.xls’ files by default, which is the Native Format of Excel.

F. **“OCR”** refers to optical character recognition, or the result of the process by which a hard copy or non-searchable Electronic Document is analyzed by a computer in order to glean the textual content of the document.

2. Production Format

With the exception of documents to be provided in Native Format described in Paragraph E, documents shall be produced as Bates Numbered tagged image file format (“TIFF” or “JPEG”)

images accompanied by an image Load File, a data Load File with fielded metadata, document-level Extracted Text for ESI, and OCR text for scanned hard copy documents and ESI that does not contain extractable text.

A. **De-duplication.** The producing party shall use industry standard Md5, SHA1, or SHA256 hash comparison method to globally de-duplicate all files identified for production. Loose Electronic Documents shall not be compared to email attachments for deduplication purposes. Emails shall be deduplicated at the family level (i.e. an email and its attachments shall be considered together for deduplication purposes). Hard copy documents containing handwritten notes shall not be considered as duplicative of any other document.

B. **ESI Processing Time zone.** All ESI shall be processed in the industry standard time zone, Coordinated Universal Time (UTC).

C. **Document Unitization.** Where documents with attachments are produced, they shall be attached in the same manner as included in the original file. Where documents are produced and all attachments thereto are not included, the producing party shall identify the missing attachments by means of a “place holder” file and explain the reason for their non-production. Documents that are segregated or separated from other documents, whether by inclusion of binders, files, dividers, tabs, clips or any other method, shall be produced in a manner that reflects these divisions. Where the producing party converts paper documents into electronic format, distinct documents must not be merged into a single record, and single documents must not be split into multiple records (*i.e.*, paper documents must be logically unitized). In the case of an organized compilation of separate documents – for example, a binder containing several separate documents behind numbered tabs – the document behind each tab should be scanned separately.

D. **Production Delivery.** Productions shall be delivered via FTP, or another secure data transfer method.

E. **JPEG/TIFF Image Requirements.**

i. JPEG images shall be produced when color is present in an image. TIFF images preferable for black and white images. JPEG/TIFF images shall be provided in single-page format and consecutively Bates Numbered.

ii. Images shall include the following content where present:

a. For word processing files (*e.g.*, Microsoft Word): Comments, “tracked changes,” similar in-line editing and all hidden content contained within the document.

b. For presentation files (*e.g.*, Microsoft PowerPoint): Speaker notes, comments and all other hidden content.

c. For spreadsheet files (*e.g.*, Microsoft Excel – if applicable): Hidden columns, rows, and sheets; comments, and “tracked changes” and any similar in-line editing or hidden content.

F. **Native Production Requirements.**

i. Spreadsheet files (*e.g.*, Microsoft Excel and .Csv files) are to be provided in Native Format.

a. In lieu of a full TIFF image version of each spreadsheet file, a single placeholder image bearing the relevant Bates number and confidentiality designation shall be produced.

b. When redaction of a spreadsheet is necessary, native redaction using an industry standard software solution (*i.e.* Mylili Blackout) is preferred. If native redaction is not possible, a redacted full TIFF version may be produced provided that the spreadsheet is manually formatted for optimal printing. If the spreadsheet requiring redaction is not reasonably useable with native redaction or in TIFF format, the Parties shall meet-and-confer to determine a suitable production format.

c. If redactions within a native spreadsheet are necessary, contact the receiving party to meet-and-confer to discuss the process, and provide a means to identify such documents in the production.

ii. Media files (e.g. .mp3, .wmv, etc.) shall be produced in Native Format with a single placeholder image bearing the relevant Bates number and confidentiality designations.

iii. Contact the receiving party to discuss a suitable production format for any proprietary or non-standard file types that require special software or technical knowledge for review, Databases or Database reports, or any data not covered by the technical specifications in this protocol. This includes data collected from collaboration and chat applications such as Slack, Microsoft Teams, Jira, mobile device data, etc.

iv. Any responsive files that cannot be accurately rendered in a reviewable TIFF format shall be produced in Native Format.

G. Load File Requirements.

i. The parties will provide a Concordance compatible data Load File with each production volume using standard ASCII delimiters and Unicode containing a header row listing the metadata fields included in the production volume.

ii. Image Load Files shall be produced in Concordance/Opticon compatible format.

H. Extracted Text/OCR Requirements.

i. Electronically extracted text must be provided for documents collected from electronic sources. Text generated via OCR must be provided when possible for all documents that do not contain electronically extractable text (e.g. non-searchable PDF files or JPG images) and for redacted and hard copy documents. The parties will not degrade the searchability of documents as part of the document production process.

ii. Document text shall be provided as separate, document-level text files and not be embedded in the metadata Load File.

iii. Text files shall be named according to the beginning Bates number of the document to which they correspond.

iv. If a document is provided in Native Format, the text file shall contain the Extracted Text of the native file.

v. A path to each Extracted Text file on the delivery media shall be included in a Load File field, or in a separate cross-reference file.

I. **Metadata.** The Parties will produce the following metadata fields, where available:

i. **BegBates** -- Beginning Bates number.

ii. **EndBates** -- Ending Bates number.

iii. **BegAttach** -- Bates number of the first page of a family range.

iv. **EndAttach** -- Bates number of the last page of a family range.

v. **PageCount** -- Number of pages in a Document.

vi. **FileExtension** -- Original file extension as the Document was maintained in the ordinary course.

vii. **FileSize** -- File size in bytes.

viii. **DocTitle** -- Document title as stored in file metadata.

ix. **DocSubject** -- Any value populated in the Subject field of the Document properties.

x. **Custodian** -- Primary custodian full name.

xi. **AllCustodians/DeDupe Custodians** – All custodians from whom a document was collected.

xii. **Author** -- Document author information for non-email.

xiii. **Email From**

xiv. **Email To**

xv. **Email CC**

xvi. **Email BCC**

xvii. **Email Subject**

xviii. **Attachments** -- Name of attached file(s) as maintained in the ordinary course of business.

xix. **DateCreated** -- File date and time created MM/DD/YYYY HH:MM AM/PM.

xx. **DateModified** -- File date and time modified MM/DD/YYYY HH:MM AM/PM.

xxi. **DateLastSaved** – File date and time last saved MM/DD/YYYY HH:MM AM/PM.

xxii. **DateSent** -- Email date and time sent MM/DD/YYYY HH:MM AM/PM.

xxiii. **DateReceived** -- Email date and time received. MM/DD/YYYY HH:MM AM/PM.

xxiv. **FileName** -- Name of the file as maintained in the ordinary course of business with extension.

xxv. **MD5Hash** -- The computer-generated MD5 Hash value for each Document.

xxvi. **TextPath** -- The path to the corresponding text file for each record on the delivery media, including filename.

xxvii. **NativePath** -- The path to the native-format file corresponding to the record on the delivery media, including the file name (if a native-format file is provided).

3. **Documents Protected from Discovery**

A. Documents that contain both privileged and non-privileged information shall be produced with the privileged information redacted in such a way as to show the location of the redaction within the Document.